



Fluctuation Results for General Block Spin Ising Models

Holger Knöpfel¹ · Matthias Löwe¹ · Kristina Schubert²  · Arthur Sinulis³

Received: 8 March 2019 / Accepted: 13 January 2020 / Published online: 1 February 2020
© The Author(s) 2020

Abstract

We study a block spin mean-field Ising model, i.e. a model of spins in which the vertices are divided into a finite number of blocks with each block having a fixed proportion of vertices, and where pair interactions are given according to their blocks. For the vector of block magnetizations we prove Large Deviation Principles and Central Limit Theorems under general assumptions for the block interaction matrix. Using the exchangeable pair approach of Stein's method we establish a rate of convergence in the Central Limit Theorem for the block magnetization vector in the high temperature regime.

Keywords Block spin Ising models · Central limit theorem · Large deviation principle · Phase transition · Stein's method

Mathematics Subject Classification Primary 60F05 · 60F10 · Secondary 82B20

1 Introduction

Mean-field block models were introduced as an approximation of a lattice model of a meta-magnet, see e.g. formula (4.1) in [24]. Furthermore, they can arise in disordered systems with random pair interactions, studied for example in [9,31,32]. Later, they were rediscovered as interesting models for statistical mechanics systems, see [8,17,19,25,27], as well as models

Communicated by Hal Tasaki.

✉ Kristina Schubert
kristina.schubert@tu-dortmund.de
Holger Knöpfel
Holger.Knoepfel@ruhr-uni-bochum.de
Matthias Löwe
malowe@math.uni-muenster.de
Arthur Sinulis
asinulis@math.uni-bielefeld.de

¹ Fachbereich Mathematik und Informatik, Universität Münster, Einsteinstraße 62, 48149 Münster, Germany

² Fakultät für Mathematik, TU Dortmund, Vogelpothsweg 87, 44227 Dortmund, Germany

³ Fakultät für Mathematik, Universität Bielefeld, Postfach 100131, 33501 Bielefeld, Germany

for social interactions between several groups, e.g. in [1,20,29]. This latter approach follows very much the social re-interpretation for *one* group of the Curie–Weiss model in [6] or of the Hopfield model in [10] or [26]. A third source of interest in mean-field spin block models is a statistical point of view. In [3], the authors gave another analysis of the bipartite mean-field Ising block model with equal block sizes, and asked the question whether one can recover the blocks from several observations from this model, and if so, how many observations are needed. In this aspect, the block spin models are related to the stochastic block models from random graph theory. These have been in the center of interest in statistics and probability theory over the past couple of years (see, e.g. [2,21]). The statistical interest in them arises from their relation to graphical models. In this framework a major question is always how to reconstruct the block structure under sparsity assumptions (see e.g. [4,5,28]).

Our starting point is [27]. There, the fluctuations of an order parameter for a two-groups block model with equal block sizes were analyzed on the level of large deviations principles (LDPs, for short) and central limit theorems (CLTs). Starting from these results, there are several natural questions. First: Can these results be also proven for systems with not necessarily identical block sizes? Second: Can we generalize our results to the situation of more than two groups? And third: Can we give a speed of convergence for the CLT? The main goal of the current note is to (partially) answer these questions. To this end, we will present a new approach to mean-field block spin models, via the corresponding block interaction matrix. Moreover, to obtain a speed of convergence in the CLT, we will employ Stein’s method as in [7,14] for the standard mean-field Ising, or Curie–Weiss model.

The rest of this note is organized in the following way. In the remaining part of this introduction, we define our model in a way that makes it accessible to our techniques in Sects. 2 and 3, and state our main results. Section 2 is devoted to the proof of the LDP results. Afterwards, we analyze the critical points of the rate function and obtain the *mean field equations*, showing that in the high temperature case the only maximum is 0, whereas in the low temperature case there are nonzero maximizers, and we obtain a solution for a special class of block interaction matrices. In Sect. 3 we prove the CLT for the order parameter of the model in two ways. One uses the classical Hubbard–Stratonovich transformation. This was already used for proving the CLT for the magnetization in the Curie–Weiss model in [16], and also is the core technique for the CLT in [27]. The second proof uses a multivariate version of the exchangeable pair approach in Stein’s method, developed in [30]. Lastly, Sect. 4 contains a discussion of some of the results and further open questions.

1.1 The Model

The block spin Ising model will be characterized by two quantities, a number $k \in \mathbb{N}$ —*number of blocks*—and a symmetric, positive definite matrix $A \in \mathbb{R}^{k \times k}$, which is the *block interaction matrix*. A_{ij} will determine the strength of interaction between two particles in block i and j respectively. Here, $\mathbb{R}^{r_1 \times r_2}$ is the set of all r_1 by r_2 matrices with real entries.

Let $N(n)$ be a strictly increasing subsequence of \mathbb{N} . For a system of size $N = N(n)$ let $B_1^{(n)}, \dots, B_k^{(n)} \subset \{1, \dots, N\}$ be a partition of $\{1, \dots, N\}$ into k blocks. Without loss of generality, we assume that the indices in the blocks are ordered, i.e. if $i_0 \in B_i^{(n)}$ and $j_0 \in B_j^{(n)}$ and $i < j$, it follows $i_0 < j_0$. We call $|B_i^{(n)}|$ the *block size* of the i th block. Note that, in particular, we have a system of size N , where for $n \in \mathbb{N}$

$$N = N(n) = \sum_{i=1}^k |B_i^{(n)}|.$$

Define for each $n \in \mathbb{N}$ the matrix of the relative block sizes

$$\Gamma_n := \text{diag} \left(\frac{\sqrt{|B_1^{(n)}|}}{\sqrt{N}}, \dots, \frac{\sqrt{|B_k^{(n)}|}}{\sqrt{N}} \right) \in \mathbb{R}^{k \times k}.$$

We assume that for each $i = 1, \dots, k$ the limit

$$\gamma_i := \lim_{n \rightarrow \infty} \sqrt{\frac{|B_i^{(n)}|}{N}} \in (0, 1)$$

exists, so that the matrix of asymptotic relative block sizes

$$\Gamma_\infty := \text{diag}(\gamma_1, \dots, \gamma_k) \in \mathbb{R}^{k \times k}$$

is invertible. If the k partition blocks are asymptotically of the same size, i.e.

$$\Gamma_\infty = \frac{1}{\sqrt{k}} \text{Id} \quad \text{resp.} \quad \gamma_i = \frac{1}{\sqrt{k}} \text{ for } i = 1, \dots, k,$$

we call this the *uniform case*. The block spin Ising model with k blocks of sizes $|B_1^{(n)}|, \dots, |B_k^{(n)}|$ and block interaction matrix A is defined as the Ising model with interaction matrix

$$J_n := \frac{1}{N} \begin{pmatrix} A_{11} O(|B_1^{(n)}|, |B_1^{(n)}|) & \cdots & A_{1k} O(|B_1^{(n)}|, |B_k^{(n)}|) \\ \vdots & & \vdots \\ A_{k,1} O(|B_k^{(n)}|, |B_1^{(n)}|) & \cdots & A_{kk} O(|B_k^{(n)}|, |B_k^{(n)}|) \end{pmatrix},$$

where $O(m, n) \in \mathbb{R}^{m \times n}$ is the matrix with all entries equal to 1. We denote this model by μ_{J_n} . More precisely, μ_{J_n} is the probability measure on $\{-1, +1\}^N$, $N = N(n)$, defined by

$$\mu_{J_n}(x) = Z_n^{-1} \exp(H_n(x)) = Z_n^{-1} \exp\left(\frac{1}{2} \langle x, J_n x \rangle\right) = Z_n^{-1} \exp\left(\frac{1}{2} \sum_{i,j=1}^N (J_n)_{ij} x_i x_j\right).$$

Here, of course, Z_n is the partition function

$$Z_n := \sum_{x \in \{-1, +1\}^N} \exp\left(\frac{1}{2} \sum_{i,j=1}^N (J_n)_{ij} x_i x_j\right).$$

Note that, contrary to the usual convention, we do not require the diagonal of J_n to be zero for technical convenience. However, since $x_i^2 = 1$, both J_n and its “dediagonalized” version $\tilde{J}_n = J_n - \text{diag}(J_{ii})$ give rise to the same Ising model. Here and in the sequel, $\text{diag}(\lambda_1, \dots, \lambda_l)$ is a diagonal $l \times l$ matrix with values $\lambda_1, \dots, \lambda_l$ on its diagonal. Lastly, for any $p, q \in [1, \infty]$ and any matrix $A \in \mathbb{R}^{k \times k}$ we define the operator norm

$$\|A\|_{p \rightarrow q} := \sup_{x \in \mathbb{R}^k : \|x\|_p = 1} \|Ax\|_q.$$

1.2 Main Results

We prove results on the fluctuations of the block magnetization vector on different scales. In what follows, we use the non-normalized and normalized versions of the block magnetization vector defined as

$$\begin{aligned} m^{(n)} &= m^{(n)}(x) = (m_1^{(n)}(x), \dots, m_k^{(n)}(x)) = \left(\sum_{j \in B_i^{(n)}} x_j \right)_{i=1, \dots, k}, \\ \tilde{m}^{(n)} &= \tilde{m}^{(n)}(x) = (\tilde{m}_1^{(n)}(x), \dots, \tilde{m}_k^{(n)}(x)) = \left(\frac{1}{|B_i^{(n)}|} \sum_{j \in B_i^{(n)}} x_j \right)_{i=1, \dots, k}, \\ \widehat{m}^{(n)} &= \widehat{m}^{(n)}(x) = (\widehat{m}_1^{(n)}(x), \dots, \widehat{m}_k^{(n)}(x)) = \left(\frac{1}{\sqrt{|B_i^{(n)}|}} \sum_{j \in B_i^{(n)}} x_j \right)_{i=1, \dots, k}. \end{aligned}$$

Note that this allows us to rewrite the Hamiltonian H_n of μ_{J_n} as

$$H_n(x) = \frac{1}{2N} \langle m^{(n)}, A m^{(n)} \rangle = \frac{1}{2} \langle \widehat{m}^{(n)}, \Gamma_n A \Gamma_n \widehat{m}^{(n)} \rangle = \frac{N}{2} \langle \Gamma_n^2 A \Gamma_n^2 \tilde{m}^{(n)}, \tilde{m}^{(n)} \rangle,$$

which we use tacitly.

We begin by presenting the large deviation results. The first result is a generalization of [27, Theorem 2.1]. In that paper, an LDP for $\tilde{m}^{(n)}$ was proved in the situation of $k = 2$ blocks of equal size. Here we analyze the general case.

Theorem 1 *Let $k \in \mathbb{N}$ and A be a block interaction matrix. The sequence $(\tilde{m}^{(n)})_{n \in \mathbb{N}}$ satisfies an LDP under $(\mu_{J_n})_{n \in \mathbb{N}}$ with speed N and rate function*

$$J(x) := \sup_{y \in \mathbb{R}^k} I(y) - I(x),$$

where

$$I(x) := \frac{1}{2} \langle x, \Gamma_\infty^2 A \Gamma_\infty^2 x \rangle - \sum_{i=1}^k \gamma_i^2 L^*(x_i),$$

and L^* denotes the convex conjugate of $\log \cosh$, i.e.

$$L^*(x) := \frac{1}{2} (1+x) \log(1+x) + \frac{1}{2} (1-x) \log(1-x) \quad x \in [-1, +1].$$

More precisely, in the notion of large deviations, the sequence of push-forwards $(\tilde{m}^{(n)} \circ \mu_{J_n})_{n \in \mathbb{N}}$ satisfies an LDP with speed N and the rate function I .

In the special case of asymptotically uniform block sizes the function I is related to the matrix A in an even more straightforward way, since in this case

$$I(x) = \frac{1}{2k^2} \langle x, Ax \rangle - \frac{1}{k} \sum_{i=1}^k L^*(x_i).$$

We show that the rate function I has a unique minimum at 0 in the case $\|\Gamma_\infty^2 A \Gamma_\infty^2\|_{2 \rightarrow 2} \leq 1$, which yields the following corollary.

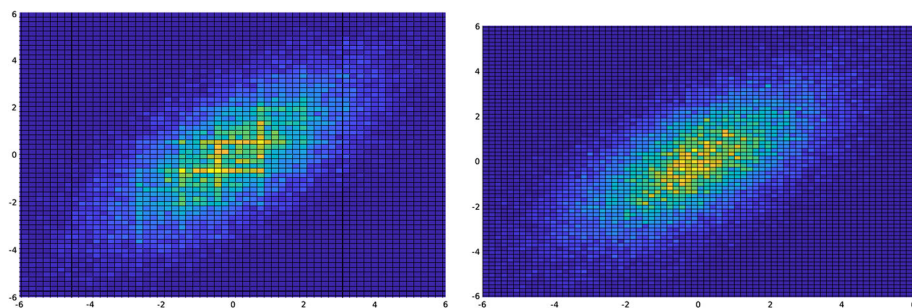


Fig. 1 A visualization of the block magnetization vector $\tilde{m}^{(n)}$ (left) for $n = 500$, using the Glauber dynamic for sampling, and a heat map for the limiting normal distribution. Here, we choose $k = 2$, $A = \begin{pmatrix} 1.1 & 0.6 \\ 0.6 & 1.1 \end{pmatrix}$ and the uniform case

Corollary 1 *Under the general assumptions, if $\|\Gamma_\infty A \Gamma_\infty\|_{2 \rightarrow 2} \leq 1$, the normalized vector of magnetizations $\tilde{m}^{(n)}$ converges to 0 exponentially fast in μ_{J_n} -probability. By this we mean more precisely, for each $\varepsilon > 0$ there is a constant I_ε such that*

$$\mu_{J_n}(\|\tilde{m}^{(n)}\|_2 \geq \varepsilon) \leq \exp(-N I_\varepsilon).$$

Let us discuss the large deviation results. In the classical Curie–Weiss model, i.e. the case $k = 1$, there is a phase transition: The limiting behavior of $\tilde{m}^{(n)}$ changes, depending on whether $A_{11} \leq 1$ (the high temperature regime), or $A_{11} > 1$ (the low temperature regime) (see [15] for an extensive treatment of this model). A corresponding phase transition can be observed in our model. This is stated in [18] for the bipartite model. In [25] the authors prove the existence of such a phase transition using the method of moments. Of course, with that method one cannot obtain an exponential speed of convergence as in Corollary 1. In accordance with the notion in the classical Curie–Weiss model, we will call these different parameter regimes the *high temperature* and *low temperature* regime, respectively. Here, the high temperature regime corresponds to $\|\Gamma_\infty A \Gamma_\infty\|_{2 \rightarrow 2} \leq 1$ and the low temperature regime to $\|\Gamma_\infty A \Gamma_\infty\|_{2 \rightarrow 2} > 1$. In the special case of asymptotically uniform block sizes (i.e. $\Gamma_\infty = \frac{1}{\sqrt{k}} \text{Id}$) these conditions reduce to $\|A\|_{2 \rightarrow 2} \leq k$ and $\|A\|_{2 \rightarrow 2} > k$ respectively.

Next, we consider the scaled block magnetization vector $\hat{m}^{(n)}$. Again, in the classical (i.e. one-dimensional) case it is known that the magnetization satisfies a central limit theorem with variance $\sigma^2 = (1 - A_{11})^{-1}$ whenever $A_{11} < 1$. The following theorem is a generalization of this phenomenon (Fig. 1).

Theorem 2 *Let $k \in \mathbb{N}$ and A be a block interaction matrix. In the high temperature regime we have*

$$\hat{m}^{(n)} \Rightarrow \mathcal{N}(0, \Sigma_\infty) = \mathcal{N}\left(0, (\text{Id} - \Gamma_\infty A \Gamma_\infty)^{-1}\right).$$

Consequently, in the uniform case

$$\hat{m}^{(n)} \Rightarrow \mathcal{N}\left(0, \left(\text{Id} - \frac{1}{k} A\right)^{-1}\right).$$

Note that Σ_∞ exists, and it can be expanded into a Neumann series. Moreover, if $\Gamma_\infty A \Gamma_\infty = V^T \Lambda V$ is an orthogonal decomposition, then $\Sigma_\infty = V^T \text{diag}((1 - \lambda_i)^{-1}) V$. Again, a similar statement is derived in [25] using the method of moments.

Furthermore, we can treat the critical case. In the Curie–Weiss model, for $\beta = 1$, the quantity $N^{-3/4} \sum_{i=1}^N \sigma_i$ converges weakly to a measure with Lebesgue-density $g_1(x) := Z^{-1} \exp\left(-\frac{x^4}{12}\right)$ (see e.g. [15, Theorem V.9.5]). As proven in [27] and [18] a similar statement holds true for the vector of magnetizations in the case of $k = 2$ blocks. The next theorem gives a further generalization of this fact in the case $k \geq 2$. Moreover, it shows that statistics associated to the orthogonal decomposition of the block interaction matrix give rise to k asymptotically independent random variables with either a Gaussian distribution or a distribution with a Lebesgue-density g_1 .

In the multidimensional critical case $\|\Gamma_\infty A \Gamma_\infty\|_{2 \rightarrow 2} = 1$ we restrict to the uniform case with a simple eigenvalue $\lambda_k = k$, i.e. we have $A = V^T \text{diag}(\lambda_1, \dots, \lambda_{k-1}, k) V$. Let $\Gamma_n A \Gamma_n = V_n^T \Lambda_n V_n$, be the orthogonal decomposition, where V_n is a unitary $k \times k$ -matrix and Λ_n a diagonal $k \times k$ -matrix. If we define the normalized vector

$$w' = w'_n := \text{diag}(N^{-1/2}, \dots, N^{-1/2}, N^{-3/4}) V m^{(n)}$$

and the matrix

$$\hat{C}_N := \text{diag}(\lambda_1, \dots, \lambda_{k-1}, kN^{1/2}),$$

we have the following result.

Theorem 3 *Under the above assumptions let $Y_n \sim \mathcal{N}(0, \hat{C}_N^{-1})$ and $X_n \sim \mu_{J_n}$ be independent random variables, defined on a common probability space. Then $w'_n(X_n) + Y_n$ converges in distribution to a probability measure with density*

$$\tilde{g}_k(x) := \tilde{Z}^{-1} \exp\left(-\frac{1}{2} \sum_{i=1}^{k-1} \left(\lambda_i - \frac{\lambda_i^2}{k}\right) x_i^2 - \frac{k^3}{12} x_k^4 \sum_{i=1}^k V_{ki}^4\right) \quad (1)$$

for a suitable normalization \tilde{Z} that makes the expression (1) a probability density.

Thus, the vector $(w'_n(X_n)_j)_{j=1, \dots, k-1}$ converges to a normal distribution with covariance matrix $\Sigma = \text{diag}((k - \lambda_j)^{-1})$ and the random variable $w'_n(X_n)_k$ converges to a distribution with Lebesgue-density $Z^{-1} \exp\left(-\left(\frac{k^3}{12} \sum_{i=1}^k V_{ki}^4\right) x^4\right) dx$.

We believe it is possible to extend Theorem 3 to the case where the eigenvalue k has multiplicity greater than 1, by appropriately rescaling all the eigenvectors which belong to the eigenvalue k .

Note that the parameter $\sigma^2 := k^3/12 \sum_{i=1}^k V_{ki}^4$ is directly related to the variance of a random variable with that distribution; indeed, a short calculation shows that for $X \sim \exp(-\sigma^2 x^4) dx$ we have $\text{Var}(X) = c\sigma^{-1}$, where c is an absolute constant. Moreover, $\sum_{i=1}^k V_{ki}^4 = \|v_k\|_4^4$, where v_k is the eigenvector belonging to the eigenvalue k .

In a final step, we establish convergence rates in the CLT in the high temperature case for a special class of functions. We use the exchangeable pair approach of Stein's method, that was also used in [14] and [7] in the case of the Curie–Weiss model. The proof of the next result will rely on a multivariate version of Stein's method proven in [30]. To this end, define the function class

$$\mathcal{F}_3 := \left\{ h : \mathbb{R}^k \rightarrow \mathbb{R} : h \in C^3(\mathbb{R}^k), \max_{j=1,2,3} \max_{\alpha=(\alpha_j)_j} \sup_{x \in \mathbb{R}^k} |\partial^\alpha h|(x) \leq 1 \right\}$$

of all three times differentiable functions with all partial derivatives (up to order three) bounded.

Theorem 4 Assume that $\|\Gamma_\infty A \Gamma_\infty\|_{2 \rightarrow 2} < 1$ and for each $n \in \mathbb{N}$ let $\Sigma_n := \mathbb{E} \hat{m}^{(n)} (\hat{m}^{(n)})^T$. For $Z \sim \mathcal{N}(0, \text{Id})$, we have

$$\sup_{h \in \mathcal{F}_3} |\mathbb{E}_{\mu_{J_n}} (h(\hat{m}^{(n)})) - \mathbb{E} h(\Sigma_n^{1/2} Z)| = O(N^{-1/2}).$$

2 Proofs of the Large Deviation Results and the Mean-Field Equations

Let us start off by proving the LDP result for the rescaled block magnetization vector $\tilde{m}^{(n)}$. Recall the notion of an LDP (for which we also refer to [13] and [12]): If \mathcal{X} is a Polish space and $(a_n)_{n \in \mathbb{N}}$ is an increasing sequence of non-negative real numbers, we say that a sequence of probability measures $(\nu_n)_n$ on \mathcal{X} satisfies a *large deviation principle* with speed a_n and rate function $I : \mathcal{X} \rightarrow \mathbb{R}$ (i.e. a lower semi-continuous function with compact level sets $\{x : I(x) \leq L\}$ for all $L > 0$), if for all Borel sets $B \in \mathcal{B}(\mathcal{X})$ we have

$$-\inf_{x \in \text{int}(B)} I(x) \leq \liminf_{n \rightarrow \infty} \frac{\log \nu_n(B)}{a_n} \leq \limsup_{n \rightarrow \infty} \frac{\log \nu_n(B)}{a_n} \leq -\inf_{x \in \text{cl}(B)} I(x),$$

where $\text{int}(B)$ and $\text{cl}(B)$ denote the topological interior and closure of a set B , respectively.

We say that a sequence of random variables $X_n : \Omega \rightarrow \mathcal{X}$ satisfies an LDP with speed a_n and rate function $I : \mathcal{X} \rightarrow \mathbb{R}$ under a sequence of measures μ_n if the push-forward sequence $\nu_n := \mu_n \circ X_n$ satisfies an LDP with speed a_n and rate function I .

To prove Theorem 1, we will need the following lemma.

Lemma 1 Let \mathcal{X} be a Polish space and assume that a sequence of measures $(\mu_n)_{n \in \mathbb{N}}$ on \mathcal{X} satisfies an LDP with speed n and rate function I . Let $F : \mathcal{X} \rightarrow \mathbb{R}$ be a continuous function which is bounded from above and $\eta_n : \mathcal{X} \rightarrow \mathbb{R}$ a sequence of functions such that $\|\eta_n\|_{L^\infty(\mu_n)} \rightarrow 0$. Then the sequence of measures

$$d\tilde{\mu}_n = \exp(nF + n\eta_n) d\mu_n$$

satisfies an LDP with speed n and rate function

$$J(x) = \sup_{\lambda \in \mathcal{X}} (F(\lambda) - I(\lambda)) - (F(x) - I(x)).$$

Proof Note that this is a slight modification of the *tilted LDP*, which is an immediate consequence of Varadhan's Lemma ([13, Theorem III.17]). Indeed, according to this tilted LDP, the sequence of measures $(\nu_n)_n$ with μ_n -density $\exp(nF)$ satisfies an LDP with speed n and rate function J . Since for any $n \in \mathbb{N}$ and any $B \in \mathcal{B}(\mathcal{X})$ the inequalities

$$e^{-2n\|\eta_n\|_{L^\infty(\mu_n)}} \nu_n(B) \leq \tilde{\mu}_n(B) \leq e^{2n\|\eta_n\|_{L^\infty(\mu_n)}} \nu_n(B)$$

hold, this easily implies an LDP for $(\tilde{\mu}_n)_n$ with speed n and the same rate function J due to $\|\eta_n\|_{L^\infty(\mu_n)} \rightarrow 0$. \square

Proof of Theorem 1 First, note that under the uniform measure μ_0 (i.e. $A \equiv 0$) we have

$$\mathbb{E}_{\mu_0} \exp \left(N \left\langle t, \tilde{m}^{(n)} \right\rangle \right) = \prod_{i=1}^k \cosh \left(t_i \frac{N}{|B_i^{(n)}|} \right)^{|B_i^{(n)}|},$$

so that

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E}_{\mu_0} \exp \left(N \left\langle t, \tilde{m}^{(n)} \right\rangle \right) = \sum_{i=1}^k \gamma_i^2 \log \cosh \left(\frac{t_i}{\gamma_i^2} \right).$$

By the Gärtner-Ellis Theorem ([12, Theorem 2.3.6]), $\tilde{m}^{(n)}$ satisfies an LDP under μ_0 with speed N and rate function

$$J_{\mu_0}(x) := \sup_{t \in \mathbb{R}^k} \left(\langle t, x \rangle - \sum_{i=1}^k \gamma_i^2 \log \cosh \left(\frac{t_i}{\gamma_i^2} \right) \right) = \sum_{i=1}^k \gamma_i^2 L^*(x_i),$$

where $L^*(x)$ is the convex conjugate of $\log \cosh$. Next, it is easy to see that we can rewrite the μ_0 -density of μ_{J_n} as

$$\frac{d\mu_{J_n}}{d\mu_0}(x) = \exp \left(\frac{N}{2} \left\langle (\Gamma_n^2 A \Gamma_n^2) \tilde{m}^{(n)}, \tilde{m}^{(n)} \right\rangle \right) = \exp \left(N F(\tilde{m}^{(n)}) - N \eta_n(\tilde{m}^{(n)}) \right),$$

where

$$F(x) = \frac{1}{2} \langle \Gamma_\infty^2 A \Gamma_\infty^2 x, x \rangle = \frac{1}{2} \langle \Gamma_\infty^2 A \Gamma_\infty^2 x, x \rangle \wedge \frac{1}{2} k \|\Gamma_\infty^2 A \Gamma_\infty^2\|_{2 \rightarrow 2},$$

$$\eta_n(x) = \frac{1}{2} \langle \Gamma_\infty^2 A \Gamma_\infty^2 x, x \rangle - \frac{1}{2} \langle \Gamma_n^2 A \Gamma_n^2 x, x \rangle.$$

Note that we artificially inserted the truncation in F to emphasize the boundedness of $F(\tilde{m}^{(n)})$. This does not affect the quadratic form, as

$$\left| \frac{1}{2} \langle \Gamma_\infty^2 A \Gamma_\infty^2 \tilde{m}^{(n)}, \tilde{m}^{(n)} \rangle \right| \leq \frac{1}{2} \|\Gamma_\infty^2 A \Gamma_\infty^2\|_{2 \rightarrow 2} \|\tilde{m}^{(n)}\|_2^2 \leq \frac{k}{2} \|\Gamma_\infty^2 A \Gamma_\infty^2\|_{2 \rightarrow 2}.$$

Moreover, F is obviously continuous and η_n satisfies

$$\|\eta_n\|_\infty \leq k \|\Gamma_\infty^2 A \Gamma_\infty^2 - \Gamma_n^2 A \Gamma_n^2\| \rightarrow 0$$

on the support of $\mu_0 \circ \tilde{m} = [-1, 1]^k$, so that the assertion follows from Lemma 1. \square

2.1 The Mean-Field Equations

Theorem 1 states that the function $I(x) = \frac{1}{2} \langle x, \Gamma_\infty^2 A \Gamma_\infty^2 x \rangle - \sum_{i=1}^k \gamma_i^2 L^*(x_i)$ determines the asymptotic behavior of the magnetization, and thus the critical points of I are of utter importance. These satisfy the so-called *mean-field equations*

$$\begin{aligned} x_1 &= \tanh((A \Gamma_\infty^2 x)_1) = \tanh \left(\sum_{j=1}^k A_{1j} \gamma_j^2 x_j \right) \\ &\vdots \\ x_k &= \tanh((A \Gamma_\infty^2 x)_k) = \tanh \left(\sum_{j=1}^k A_{kj} \gamma_j^2 x_j \right). \end{aligned} \tag{2}$$

For example, in the well-studied case $k = 2$, choosing

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{12} & A_{22} \end{pmatrix} \quad \text{and} \quad \Gamma_\infty^2 = \begin{pmatrix} \gamma & 0 \\ 0 & 1 - \gamma \end{pmatrix}$$

for a positive definite matrix A and $\gamma \in (0, 1)$ Eq. (2) reduce to

$$\begin{aligned} x_1 &= \tanh(\gamma A_{11} x_1 + (1 - \gamma) A_{12} x_2), \\ x_2 &= \tanh(\gamma A_{12} x_1 + (1 - \gamma) A_{22} x_2). \end{aligned}$$

Whereas for the two-dimensional fixed point problem the existence of a solution can be shown by monotonicity arguments, the existence of a solution to (2) for general k is more involved. First off, we show that in the high temperature regime the only critical point of I is 0. This will immediately yield Corollary 1.

Proof of Corollary 1 In the sense of the formulation in Corollary 1, $\tilde{m}^{(n)}$ concentrates exponentially fast in the minima of the function J . However, under the condition $\|\Gamma_\infty A \Gamma_\infty\|_{2 \rightarrow 2} \leq 1$ there is only one minimum, which is zero. To see this, note that any local minimum satisfies

$$\nabla J(x) = -\Gamma_\infty^2 A \Gamma_\infty^2 x + \Gamma_\infty^2 \operatorname{artanh}(x) = 0. \quad (3)$$

Here, $\operatorname{artanh}(x)$ is understood componentwise. Clearly, 0 is a solution, and due to

$$\operatorname{Hess} I(0) = -\Gamma_\infty^2 A \Gamma_\infty^2 + \Gamma_\infty^2 = \Gamma_\infty (\operatorname{Id} - \Gamma_\infty A \Gamma_\infty) \Gamma_\infty \geq 0 \quad (4)$$

this is a local minimum. Assume there is some $y \neq 0$ solving (3), and observe that

$$\begin{aligned} \|\Gamma_\infty^2 y\|_2^2 &= \langle \Gamma_\infty^2 y, \Gamma_\infty^2 \operatorname{artanh}(y) + (\operatorname{Id} - \Gamma_\infty^2 A) \Gamma_\infty^2 y \rangle \geq \langle \Gamma_\infty^2 y, \Gamma_\infty^2 \operatorname{artanh}(y) \rangle \\ &= \sum_{i=1}^k \gamma_i^4 \operatorname{artanh}(y_i) y_i \geq \sum_{i=1}^k \gamma_i^4 y_i^2 = \|\Gamma_\infty^2 y\|_2^2. \end{aligned} \quad (5)$$

Here the first inequality follows from the general fact that the spectrum of the matrices BC and CB agree, applied to $B = \Gamma_\infty$ and $C = \Gamma_\infty A$. The last inequality follows from $\operatorname{artanh}(x)x \geq x^2$ for all $x \in (-1, 1)$, with equality for $x = 0$ only. This means that for any solution y we have equality in (5). However, equality can only hold if $y_i = 0$ whenever $\gamma_i \neq 0$. Due to our assumption $\gamma_i \in (0, 1)$, this proves the claim. \square

In contrast, in the low temperature regime, there are other solutions to the mean-field equations (2). Let us start with the following proposition showing the connection of the k -dimensional mean-field equations to the one-dimensional equations of the Curie–Weiss model. It provides an explicit formula for the solution of the k -dimensional problem in terms of the solution of the Curie–Weiss equation.

Proposition 1 Let $k \in \mathbb{N}$, $\Gamma_\infty = \frac{1}{\sqrt{k}} \operatorname{Id}$ and A be a positive semidefinite, symmetric matrix with $\|A\|_{2 \rightarrow 2} > k$. If the eigenvector v_k belonging to the largest eigenvalue λ_k can be rescaled to satisfy $v_k \in \{-1, 0, 1\}^k$, then there exists a solution $x \neq 0$ to the mean-field equations (2) and it is given by $x = m^* v_k$, where m^* is the positive one-dimensional solution of the Curie–Weiss model with temperature $\beta = \lambda_k k^{-1} > 1$.

Proof Let $m^* > 0$ be the unique positive solution of the Curie–Weiss equation $\tanh(\frac{\lambda_k}{k} x) = x$ for $\beta := \frac{\lambda_k}{k} > 1$ and define $v := m^* v_k$. We have

$$\tanh\left(\frac{1}{k} A v\right) = \tanh\left(\frac{m^*}{k} A v_k\right) = \tanh\left(\frac{m^* \lambda_k}{k} v_k\right) = \tanh\left(\frac{m^* \lambda_k}{k}\right) v_k = v,$$

where in the second-to-last step we have used explicitly that $v_k \in \{-1, 0, 1\}^k$, and so v is a critical point of I . Moreover, in this case it is easily seen that

$$\operatorname{Hess}\left(\frac{1}{2k^2} \langle x, Ax \rangle - \frac{1}{k} \sum_{i=1}^k L^*(x_i)\right)(v) = \frac{1}{k^2} A - \frac{1}{k(1 - (m^*)^2)} \operatorname{Id}$$

is negative definite. Indeed, from

$$g(x) := \frac{\operatorname{artanh}(x)}{x} - \frac{1}{1-x^2} = \sum_{k=0}^{\infty} x^{2k} \left(\frac{1}{1+2k} - 1 \right) \leq 0$$

we obtain

$$\begin{aligned} \left\langle y, \left(\frac{1}{k^2} \Lambda - \frac{1}{k(1-(m^*)^2)} \right) y \right\rangle &= \frac{1}{k} \sum_{i=1}^k y_i^2 \left(\frac{\lambda_i}{k} - \frac{1}{1-(m^*)^2} \right) \\ &\leq \frac{1}{k} \sum_{i=1}^k y_i^2 \left(\frac{\lambda_k}{k} - \frac{1}{1-(m^*)^2} \right) = \frac{g(m^*)}{k} \sum_{i=1}^k y_i^2 \\ &\leq 0. \end{aligned}$$

□

Example 1 Even though the assumptions in the previous proposition seem to be tailor-made for its proof (and the conclusion also holds true more generally), there are interesting non-trivial examples of a matrix satisfying the conditions of Proposition 1. One of them is the family of $k \times k$ matrices ($k \in \mathbb{N}$) of the form

$$A(\alpha, \beta) = (\beta - \alpha) \operatorname{Id} + \alpha O(k, k)$$

for any parameters (α, β) satisfying

$$\beta + (k-1)\alpha > k \quad \text{and} \quad \beta > \alpha. \quad (6)$$

This corresponds to k groups with an interaction parameter β within the group and α between the groups. For example, the condition (6) is satisfied whenever $\beta > \alpha > 1$.

In the general case, the conclusion of Proposition 1 holds as well. In this case the proof relies on the fact that the continuous function I has a global maximum on its (compact) domain $[-1, 1]^k$, and the next lemma excludes maxima on the boundary. Hence there is always at least one solution $y \neq 0$ (since 0 is either an inflection point or a minimum) to (2).

Lemma 2 Let I be the large deviation rate function from Theorem 1, i.e.

$$I(x) := \frac{1}{2} \langle x, \Gamma_{\infty}^2 A \Gamma_{\infty}^2 x \rangle - \sum_{i=1}^k \gamma_i^2 L^*(x_i), \quad x \in [-1, 1]^k$$

and L^* denotes the convex conjugate of $\log \cosh$.

1. I has no global maxima on the boundary of $[-1, 1]^k$.
2. If $x \in [-1, 1]^k$ satisfies the mean-field equations, we have

$$I(x) = \frac{1}{2} \sum_{i=1}^k \gamma_i^2 (x_i \operatorname{artanh}(x_i) + \log(1 - x_i^2)). \quad (7)$$

3. The set of all global maximisers has a positive distance from the boundary.

Proof (1): Assume that x is a global maximum of I on the boundary. Then there is at least one index $j \in \{1, \dots, k\}$ such that $x_j = 1$ (if $x_j = -1$, switch to $-x$ since $I(-x) = I(x)$). Rewriting the fact that x is a maximum of I , we have for any $y_j \in [-1, 1]$ and $C := \Gamma_\infty^2 A \Gamma_\infty^2$

$$\frac{1}{2} \left(\langle (\bar{x}_j, 1), C(\bar{x}_j, 1) \rangle - \langle (\bar{x}_j, y_j), C(\bar{x}_j, y_j) \rangle \right) \geq \gamma_j^2 (L^*(1) - L^*(y_j)),$$

where $\bar{x}_j \in \mathbb{R}^{k-1}$ is the vector obtained from x by deleting the j th component. If we divide both sides by $1 - y$ and let $\limsup_{y \rightarrow 1}$, the left hand side is finite, as $\frac{1}{2} \langle x, Cx \rangle \in C^\infty(\mathbb{R}^k)$, and the right hand side tends to ∞ by l'Hospital's rule. This proves statement (1).

(2): Clearly, x can only satisfy the mean-field equations if $x \in (-1, +1)^k$. For any $i = 1, \dots, k$ we have $\text{artanh}(x_i) = \sum_{j=1}^k A_{ij} \gamma_j^2 x_j = (A \Gamma_\infty^2 x)_i$. Inserting this into the function I gives

$$\begin{aligned} I(x) &= \frac{1}{2} \sum_{i=1}^k \gamma_i^2 x_i (\Gamma_\infty^2 A x)_i - \sum_{i=1}^k \gamma_i^2 L^*(x_i) = \frac{1}{2} \sum_{i=1}^k \gamma_i^2 (x_i \text{artanh}(x_i) - 2L^*(x_i)) \\ &= -\frac{1}{2} \sum_{i=1}^k \gamma_i^2 (x_i \text{artanh}(x_i) + \log(1 - x_i^2)) \\ &=: \frac{1}{2} \sum_{i=1}^k \gamma_i^2 R(x_i). \end{aligned}$$

(3): The function I is bounded in $[-1, 1]^k$, as

$$|I(x)| \leq 2 \log 2 + \frac{k}{2} \|\Gamma_\infty^2 A \Gamma_\infty^2\|_{2 \rightarrow 2}.$$

On the other hand, if there exists a sequence of maximisers approaching the boundary, i.e. for at least one i we have $x_i \rightarrow 1$, this gives $R(x_i) \rightarrow \infty$. \square

In the case of two blocks, i.e. $k = 2$, equal block sizes and the same interaction within a group, the set of maximisers of the rate function is explicitly known. Indeed, in [3, Proposition 2.1] and [27, Theorem 2.1] the authors show that for

$$A = \begin{pmatrix} \beta & \alpha \\ \alpha & \beta \end{pmatrix}$$

satisfying $\beta \geq \alpha \geq 0$ and $\beta + \alpha > 2$ (the low temperature case) the distribution of $\tilde{m}^{(n)}$ concentrates in the two points $x = (m^+((\beta + \alpha)/2), m^+((\beta + \alpha)/2))$, and $-x$. In the case $\alpha < 0$ the limit points for $\tilde{m}^{(n)}$ become $x = (m^+((\beta + \alpha)/2), -m^+((\beta + \alpha)/2))$, and $-x$. Here $m^+(b)$ is the largest solution to

$$m = \tanh(bm).$$

If $\beta + |\alpha| \leq 2$, the distribution of $\tilde{m}^{(n)}$ concentrates in the origin. For $k = 2$, we can extend this result to arbitrary block sizes.

Proposition 2 Let $k = 2$, $A = \begin{pmatrix} \beta & \alpha \\ \alpha & \beta \end{pmatrix}$ be a block interaction matrix and $\gamma_1^2 = \gamma$, $\gamma_2^2 = (1 - \gamma)$ for some $0 < \gamma < \frac{1}{2}$. In the low temperature case, if the groups are not interacting (i.e. $\alpha = 0$) there exists either two or four global maxima of I ; for $\alpha \neq 0$, there are always two global maxima of I .

Note that we have to restrict to $|\alpha| < \beta$ and $\beta > 0$ in order for A to be positive definite. Moreover, the characterization of the high temperature phase $\Gamma_\infty A \Gamma_\infty \preceq \text{Id}$ (where \preceq is the Loewner partial ordering) can be reduced to $\langle (\text{Id} - \Gamma_\infty A \Gamma_\infty) e_1, e_1 \rangle > 0$ and $\det(\text{Id} - \Gamma_\infty A \Gamma_\infty) > 0$. Thus we are in the high temperature regime if and only if

$$\beta\gamma < 1 \quad \text{and} \quad (\beta^2 - \alpha^2)\gamma(1 - \gamma) > \beta - 1.$$

Proof The case $\alpha = 0$ is an easy consequence of the statements for the one-dimensional Curie–Weiss model, since $I(x_1, x_2) = I_1(x_1) + I_2(x_2)$ and $\Gamma_\infty^2 A \Gamma_\infty^2 = \text{diag}(\beta\gamma^2, \beta(1 - \gamma^2))$.

We treat the case $\alpha > 0$ only—the case $\alpha < 0$ follows immediately from the equality $I_{\alpha, \beta}(x, y) = I_{-\alpha, \beta}(x, -y)$ (with the appropriate modifications, e.g. the maximum will be in the second quadrant instead of the first).

Due to (7) the maximum of the rate function is non-negative, let us call this maximum η . Then, $I(x, y) = \eta = 0$ implies $(x, y) = 0$, which is a contradiction to the low temperature case (recall the Hessian of I in 0 given in Eq. (4)), so that $\eta > 0$. Moreover, every global maximum (and thus local maximum, as it is not attained on the boundary) satisfies the mean-field equations, and so the value of I at any maximum is given by equation (7). As a consequence, all global maxima lie on a contour line $C_\eta := \{x \in [-1, 1]^2 : \gamma R(x_1) + (1 - \gamma)R(x_2) = 2\eta\}$, where $R(x) = x \operatorname{artanh}(x) + \log(1 - x^2)$ was defined in the previous lemma.

Firstly, let us show that in the first quadrant there can only be one such point. Due to symmetry, the global maximum will also be present in the third quadrant. For $x_1 > 0$ the points on the contour line C_η can be described by a function $x_2 = g(x_1)$, and due to the monotonicity of R the function g is non-increasing. Moreover, the solutions of the mean-field equations can be described by the functions

$$\begin{aligned} f_1(x) &:= \frac{1}{(1 - \gamma)\alpha} (\operatorname{artanh}(x) - \gamma\beta x), \\ f_2(x) &:= \frac{1}{\gamma\alpha} (\operatorname{artanh}(x) - (1 - \gamma)\beta x) \end{aligned}$$

via $x_2 = f_1(x_1)$ and $x_1 = f_2(x_2)$. The function f_1 can behave in two ways, depending on the parameter $\gamma\beta$: For $\gamma\beta \leq 1$ it increases monotonically. For $\gamma\beta > 1$ it decreases first and then increases. More precisely, in the latter case, $f_1(t) = 0$ if and only if $t \in \{0, \pm m_{\gamma\beta}\}$ for some $m_{\gamma\beta} > 0$ and f_1 is strictly increasing for $t \geq m_{\gamma\beta}$. Moreover, the curve $(x, f_1(x))$ is only in the first quadrant if $m_{\gamma\beta} < x \leq 1$. In either case, there is only one intersection point of g and f_1 in the first quadrant.

Secondly, the maximum cannot be in the second quadrant. Assume that there are solutions to the mean field equations both in the first and in the second quadrant. If we denote by m_c the zeros of $\phi_c(t) := \operatorname{artanh}(t) - ct$, for the solution in the second quadrant, we easily see that $-m_c < x < 0$ and $0 \leq y \leq m_{\beta(1-\gamma)}$. Hence

$$I(x, y) = \frac{1}{2} (\gamma R(x) + (1 - \gamma)R(y)) < \frac{1}{2} (\gamma R(m_{\gamma\beta}) + (1 - \gamma)R(m_{(1-\gamma)\beta})).$$

If there is also a solution in the first quadrant with coordinates (x^*, y^*) , we obtain analogously

$$I(x^*, y^*) > \frac{1}{2} (\gamma R(m_{\gamma\beta}) + (1 - \gamma)R(m_{(1-\gamma)\beta})).$$

This yields that the maximum must lie in the first quadrant □

Furthermore, we can treat the case $k > 2$ for uniform block sizes and special matrices. The proof is motivated by [3, Proposition 2.1].

Lemma 3 Let $k \geq 2$ and A be a block interaction matrix with positive entries such that we have for any $i = 1, \dots, k$ for two constants $c_1, c_2 > 0$ $A_{ii} = c_1$ and $\sum_{j \neq i} A_{ij} = c_2$.

In the uniform case, there are exactly two maximisers of the rate function I and they satisfy $x = m^*(1, \dots, 1)$ for m^* solving the Curie–Weiss equation $\frac{c_1+c_2}{k}x = \operatorname{artanh}(x)$.

Proof Using the equality $xy = -\frac{1}{2}(x-y)^2 + \frac{1}{2}x^2 + \frac{1}{2}y^2$ we can rewrite the rate function as

$$\begin{aligned} I(x) &= \frac{1}{k} \left(\frac{1}{2k} \sum_{i \neq j} A_{ij} x_i x_j + \frac{c_1}{2k} \langle x, x \rangle - \sum_{i=1}^k L^*(x_i) \right) \\ &= \frac{1}{k} \left(-\frac{1}{4k} \sum_{i,j} A_{ij} (x_i - x_j)^2 + \frac{c_1 + c_2}{2k} \langle x, x \rangle - \sum_{i=1}^k L^*(x_i) \right) \\ &\leq \frac{c_1 + c_2}{2k^2} \langle x, x \rangle - \frac{1}{k} \sum_{i=1}^k L^*(x_i), \end{aligned}$$

where equality only holds in the case $x_i = x_j$ for all i, j . Thus, we search for maximisers of I on the generalized diagonal $\{x \in [-1, 1]^k : x_i = x_j \forall i, j\}$. On this set we have

$$I((x, \dots, x)) = \frac{c_1 + c_2}{2k} x^2 - L^*(x),$$

i.e. it reduces to the Curie–Weiss equations in one dimension. For $c_1 + c_2 > k$ it has a unique nonzero solution m^* , and $x = m^*(1, \dots, 1)$ solves the k -dimensional maximization problem. \square

Unfortunately, the proof cannot be modified in a straightforward way to deal with non-equal block sizes, not even in the case $k = 2$. The reason is that the inequality used in the proof does not give any information on the actual maximiser in this setting (i.e. I is not maximized on any type of (weighted) diagonal). As such, we cannot reduce this to the one-dimensional setting.

Example 2 For example, Lemma 3 can be used to prove that given three positive parameters α, β, γ with $\beta > \alpha$ and $\beta + \alpha > 2\gamma$, the rate function corresponding to

$$A = \begin{pmatrix} \beta & \alpha & \gamma & \gamma \\ \alpha & \beta & \gamma & \gamma \\ \gamma & \gamma & \beta & \alpha \\ \gamma & \gamma & \alpha & \beta \end{pmatrix}$$

only has two maximisers in the uniform case. The conditions on α, β, γ ensure that A is positive definite, and it is clear that $c_1 = \beta$ and $c_2 = \alpha + 2\gamma$.

As a concluding remark let us note that the previous results imply that there is indeed a phase transition in our block spin model. However, if $k > 2$ or the block sizes are not equal, it seems hard to give a similarly explicit formula for the limit points. Nevertheless, the above observations show that there is a phase transition in a very general class of block spin models with an arbitrary number of blocks and general class of block sizes. In particular, they also justify the names “high temperature regime” and “low temperature regime”.

3 Proofs of the Limit Theorems

In this section we prove (standard and non-standard) Central Limit Theorems for the vector $\widehat{m}^{(n)}$. In the first subsection we will treat the high temperature regime. Here we derive a standard CLT using the Hubbard–Stratonovich transform. This is in spirit similar to the third section in [27] and technically related to [22]. The result can also be derived from [17], where similar techniques are used. However, the subsection also prepares nicely for Sect. 3.2, where we treat the critical case and show a non standard CLT. This generalizes results from [18] and [27]. Finally, in Sect. 3.3 we will use Stein’s method, an alternative approach to prove the CLT for $\widehat{m}^{(n)}$. This is not only interesting in its own right, but also has the advantage of providing a speed of convergence, which is missing in the case of a proof via the Hubbard–Stratonovich transform.

3.1 Central Limit Theorem: Hubbard–Stratonovich Approach

For the proof we shall use the transformed block magnetization vectors

$$\begin{aligned} w^{(n)} &:= V_n m^{(n)}, \\ \widehat{w}^{(n)} &:= V_n \widehat{m}^{(n)}, \\ \widetilde{w}^{(n)} &= V_n \Gamma_n \widetilde{m}^{(n)}, \end{aligned}$$

where $\Gamma_n A \Gamma_n = V_n^T \Lambda_n V_n$ is the orthogonal decomposition. It is easy to see that

$$H_n = \frac{1}{2N} \langle w^{(n)}, \Lambda_n w^{(n)} \rangle = \frac{1}{2} \langle \widehat{w}^{(n)}, \Lambda_n \widehat{w}^{(n)} \rangle = \frac{N}{2} \langle \Lambda_n \widetilde{w}^{(n)}, \widetilde{w}^{(n)} \rangle.$$

Proof of Theorem 2 As in [27] or [17] (both papers are inspired by [16]), we use the Hubbard–Stratonovich transform (i.e. a convolution with an independent normal distribution). For each $n \in \mathbb{N}$,

$$\mu_{J_n}(\sigma) = Z_n^{-1} \exp \left(\frac{1}{2} \langle \Lambda_n \widehat{w}^n, \widehat{w}^n \rangle \right).$$

Our first step is to prove that \widehat{w}^n converges weakly to a normal distribution. Let $Y_n \sim \mathcal{N}(0, \Lambda_n^{-1})$ be an independent sequence, which is moreover independent of $(\widehat{w}^n)_{n \in \mathbb{N}}$. We have for any $B \in \mathcal{B}(\mathbb{R}^k)$

$$\begin{aligned} \mathbb{P}(\widehat{w}^n + Y_n \in B) &= \frac{1}{Z_n} \sum_{\sigma \in \{\pm 1\}^N} \mu_{J_n}(\sigma) \int_B \exp \left(-\frac{1}{2} \langle x - \widehat{w}^n, \Lambda_n (x - \widehat{w}^n) \rangle \right) dx \\ &= \frac{2^N}{C_n Z_n} \int_B \exp \left(-\frac{1}{2} \langle x, \Lambda_n x \rangle \right) \mathbb{E}_{\mu_0} \exp \left(N \left\langle \frac{1}{\sqrt{N}} \Gamma_n V^T \Lambda_n x, \frac{1}{N} \Gamma_n^{-2} m \right\rangle \right) dx \\ &= \frac{2^N}{C_n Z_n} \int_B \exp (-\Phi_n(x)) dx, \end{aligned}$$

where we have defined

$$\begin{aligned} \Phi_n(x) &:= \frac{1}{2} \langle x, \Lambda_n x \rangle - \sum_{i=1}^k |B_i^{(n)}| \log \cosh \left(\frac{\sqrt{N}}{|B_i^{(n)}|} (\Gamma_n V_n \Lambda_n x)_i \right) \\ &= \frac{1}{2} \langle x, \Lambda_n x \rangle - \sum_{i=1}^k |B_i^{(n)}| \log \cosh \left(|B_i^{(n)}|^{-1/2} (V_n \Lambda_n x)_i \right). \end{aligned}$$

Since $\log \cosh(x) = \frac{1}{2}x^2 + O(x^4)$, we obtain

$$\begin{aligned}\Phi_n(x) &= \frac{1}{2} \langle x, A_n x \rangle - \frac{1}{2} \langle x, A_n^2 x \rangle + \frac{1}{N} O \left(\sum_{i=1}^k \frac{N}{|B_i^{(n)}|} (V_n A_n x)_i^4 \right) \\ &= \frac{1}{2} \langle x, (A_n - A_n^2) x \rangle + \frac{1}{N} O(\|\Gamma_n^{-1/2} V_n A_n x\|_4^4).\end{aligned}\quad (8)$$

For parameters $r, R > 0$ let $B_{0,r,R} := \{x \in \mathbb{R}^k : r \leq \|x\|_2^2 \leq R\}$ and decompose

$$\begin{aligned}\mathbb{P}(\widehat{w}^n + Y_n \in B) &= \frac{2^N}{C_n Z_n} \left(\int_{B \cap B_R(0)} + \int_{B \cap B_{0,R,r\sqrt{N}}} + \int_{B \cap B_{r\sqrt{N}}(0)^c} \right) \exp(-\Phi_n(x)) dx \\ &=: \frac{2^N}{C_n Z_n} (I_1 + I_2 + I_3).\end{aligned}$$

Since $A_n \rightarrow A_\infty$ (which is a consequence of the continuity of the eigenvalues) we have for any $R > 0$

$$\lim_{n \rightarrow \infty} I_1 = \int_{B \cap B_R(0)} \exp\left(-\frac{1}{2} \langle x, (A_\infty - A_\infty^2) x \rangle\right) dx.$$

Next, we will estimate (8) from below in order to obtain an upper bound for I_2 . If we define $C_{2,4} := \|\text{Id}\|_{2 \rightarrow 4}$, it follows that

$$\begin{aligned}\Phi_n(x) &\geq \frac{1}{2} \langle x, (A_n - A_n^2) x \rangle - C(r) \|\Gamma_n^{-1/2}\|_{4 \rightarrow 4} r^2 \|A_n\|_{2 \rightarrow 2}^2 \langle x, x \rangle \\ &\geq \frac{1}{2} \langle x, (A_n - A_n^2 - C(r)r^2 C) x \rangle \\ &\geq c \frac{1}{2} \langle x, x \rangle.\end{aligned}$$

Here, we have used the convergence of Γ_n to Γ_∞ to bound $\|\Gamma_n^{-1/2}\|_{4 \rightarrow 4}$ and the fact that $C(r)r^2 \rightarrow 0$ as $r \rightarrow 0$, so that the right hand side is positive definite for r small enough, uniformly in n . Thus, after taking the limit $n \rightarrow \infty$, I_2 will vanish in the limit $R \rightarrow \infty$.

Lastly, we need to show that I_3 vanishes as well. To this end, we show that we can choose $r > 0$ small enough to ensure that $\Phi_n(x) \geq \exp(-Nc)$ uniformly for $x \in B_{r\sqrt{N}}(0)^c$ and for n large enough. Since $\|A_n - A_\infty\|_{2 \rightarrow 2} \rightarrow 0$ and $\|A_\infty\|_{2 \rightarrow 2} < 1$, choose n large enough so that $\|A_n\|_{2 \rightarrow 2} < 1$ uniformly. Again, as before, it can be seen that 0 is the only minimum for n chosen that way. Indeed, after some manipulations any critical point satisfies $\Gamma_n A \Gamma_n \tanh(y) = y$, and since $\|\tanh(y)\|_2 \leq \|y\|_2$ and $\|\Gamma_n A \Gamma_n\|_{2 \rightarrow 2} < 1$, this is only possible for $y = 0$. As a consequence, for any $r > 0$ there is a constant c such that uniformly $\tilde{\Phi}_n(x) \geq c$, i.e.

$$I_3 \leq \int \mathbb{1}_{\{\|x\|_2 > r\sqrt{N}\}} \exp(-\Phi_n(x)) dx \leq \int_{B_{r\sqrt{N}}(0)^c} \exp(-N\tilde{\Phi}_n(N^{-1/2}x)) dx \rightarrow 0.$$

Lastly, choose $r > 0$ so small that $A_n - A_n^2 - C(r)r^2 C$ is uniformly positive definite, and observe that we obtain

$$\lim_{n \rightarrow \infty} \mathbb{P}(\widehat{w}^n + Y_n \in B) = \mathcal{N}(0, (A_\infty - A_\infty^2)^{-1})(B).$$

From here, it remains to undo the convolution (e.g. by using the characteristic function), giving

$$\lim_{n \rightarrow \infty} \mu_{J_n}(\widehat{w}^n \in B) = \mathcal{N}(0, (\text{Id} - \Lambda_\infty)^{-1})(B).$$

With the help of Slutsky's theorem and the definition $\widehat{m}^n = V_n^T \widehat{w}^n$ this implies

$$\mu_{J_n} \circ \widehat{m}^n \Rightarrow \mathcal{N}(0, V^T (\text{Id} - \Lambda_\infty)^{-1} V) = \mathcal{N}(0, (\text{Id} - \Gamma_\infty A \Gamma_\infty)^{-1})$$

as claimed. \square

Example 3 Consider the case $k = 2$ and

$$A_2 = \begin{pmatrix} \beta & \alpha \\ \alpha & \beta \end{pmatrix}.$$

A_2 is positive definite if $\beta \geq 0$ and $(\beta - \alpha)(\beta + \alpha) \geq 0$, i.e. if $|\alpha| \leq \beta$. We have the diagonalization

$$A_2 = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \beta + \alpha & 0 \\ 0 & \beta - \alpha \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} =: V^T \Lambda V,$$

and $w = V^T m = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} m$ corresponds to the transformation performed in [27, Theorem 1.2] (up to a factor of $\sqrt{2}$). In this case

$$\left(\text{Id} - \frac{1}{2} A_2 \right)^{-1} = \frac{2}{(\beta - 2)^2 - \alpha^2} \begin{pmatrix} 2 - \beta & \alpha \\ \alpha & 2 - \beta \end{pmatrix}$$

which is exactly the covariance matrix in [27] (again up to a factor of 2). Note that similar results have been derived in [25].

Remark 1 If $A \in M_k(\mathbb{R})$ is symmetric and positive semidefinite, then a variant of the proof shows that if we let $A = V^T \Lambda V$ with $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_l, 0, \dots, 0)$ for $l < k$, $((V\widehat{m})_i)_{i \leq l}$ converges to an l -dimensional normal distribution with covariance matrix $\Sigma_l := (\text{Id} - \Lambda_l)^{-1}$, $\Lambda_l = \text{diag}(\lambda_1, \dots, \lambda_l)$. This can be applied to the matrix A_2 above with $\alpha = \beta$, resulting in a CLT for the magnetization in a Curie–Weiss model, which of course can also be obtained by choosing $k = 1$ and $0 < \beta < 1$.

3.2 Non-central Limit Theorem

Recall the situation of Theorem 3: The block interaction matrix has eigenvalues $0 < \lambda_1 \leq \dots \leq \lambda_{k-1} < \lambda_k = k$ and we consider the uniform case, i.e. $\Gamma_\infty^2 = k^{-1}$. Moreover, we use the definitions

$$w' = \text{diag}(N^{-1/2}, \dots, N^{-1/2}, N^{-3/4}) V m^{(n)},$$

$$\widehat{C}_N = \text{diag}(\lambda_1, \dots, \lambda_{k-1}, k N^{1/2}),$$

so that

$$H_n = \frac{1}{2} \left(\widehat{C}_N w', w' \right).$$

Proof of Theorem 3 Let $Y_n \sim \mathcal{N}(0, \hat{C}_N^{-1})$ and $X_n \sim \mu_{J_n}$ be independent random variables, defined on a common probability space. We have for any Borel set $B \in \mathcal{B}(\mathbb{R}^k)$

$$\begin{aligned} \mathbb{P}(w'_n(X_n) + Y_n \in B) &= 2^N Z_n^{-1} \int_B \exp\left(-\frac{1}{2} \langle \hat{C}_N x, x \rangle\right) \mathbb{E}_{\mu_0} \exp\left(\langle x, \hat{C} w' \rangle\right) dx \\ &= \tilde{Z}_n^{-1} \int_B \exp\left(-\frac{1}{2} \langle \hat{C}_N x, x \rangle + \frac{N}{k} \sum_{i=1}^k \log \cosh((V^T \Lambda \tilde{x})_i)\right) dx \\ &= \tilde{Z}_n^{-1} \int_B \exp(-\Phi_N(x)) dx \\ &= \tilde{Z}_n^{-1} \int_B \exp\left(-N \tilde{\Phi}_N\left(\frac{x_1}{N^{1/2}}, \dots, \frac{x_{k-1}}{N^{1/2}}, \frac{x_k}{N^{1/4}}\right)\right) dx \end{aligned}$$

where we used

$$\begin{aligned} \Phi_N(x) &:= \frac{1}{2} \langle x, \hat{C}_N x \rangle - \frac{N}{k} \sum_{i=1}^k \log \cosh\left(\left(V^T \Lambda \left(\frac{x_1}{N^{1/2}}, \dots, \frac{x_{k-1}}{N^{1/2}}, \frac{x_k}{N^{1/4}}\right)\right)_i\right), \\ \tilde{\Phi}_N(x) &:= \frac{1}{2} \langle x, \Lambda x \rangle - \frac{1}{k} \sum_{i=1}^k \log \cosh\left((V^T \Lambda x)_i\right). \end{aligned}$$

Now the proof is along the same lines as the proof of the CLT in the high temperature phase, with the slight modification that we use the expansion of $\log \cosh$ to fourth order

$$\log \cosh(x) = \frac{x^2}{2} - \frac{x^4}{12} + O(x^6).$$

We again split \mathbb{R}^k into three regions, namely the inner region $I_1 = B_R(0)$ for an arbitrary $R > 0$, the intermediate region $I_2 = K_r \setminus B_R(0)$ for some arbitrary $r > 0$, where

$$K_r := \left\{x \in \mathbb{R}^k : \|(N^{-1/2}x_1, \dots, N^{-1/2}x_{k-1}, N^{-1/4}x_k)\|_\infty \leq r\right\},$$

and the outer region $I_3 := K_r^c$. Also define the rescaled vector

$$\tilde{x} := (\lambda_1 N^{-1/2}x_1, \dots, \lambda_{k-1} N^{-1/2}x_{k-1}, k N^{-1/4}x_k).$$

Firstly, in the inner region we rewrite

$$\begin{aligned} \Phi_N(x) &= \frac{1}{2} \langle x, \hat{C}_N x \rangle - \frac{N}{2k} \sum_{i=1}^k (V^T \tilde{x})_i^2 + \frac{N}{12k} \sum_{i=1}^k (V^T \tilde{x})_i^4 + \frac{N}{k} O(\|V^T \tilde{x}\|_6^6) \\ &= \frac{1}{2} \sum_{i=1}^{k-1} \left(\lambda_i - \frac{\lambda_i^2}{k}\right) x_i^2 + \frac{N}{12k} \|V^T \tilde{x}\|_4^4 + \frac{N}{k} O(\|V^T \tilde{x}\|_6^6) \\ &= \frac{1}{2} \sum_{i=1}^{k-1} \left(\lambda_i - \frac{\lambda_i^2}{k}\right) x_i^2 + \frac{k^3}{12} x_k^4 \sum_{i=1}^k V_{ki}^4 + O(N^{-1/4}) + \frac{N}{k} O(\|V^T \tilde{x}\|_6^6), \end{aligned}$$

and since the convergence of the error terms is uniform on any compact subset of \mathbb{R}^k , for any fixed $R > 0$ this yields

$$\lim_{N \rightarrow \infty} \int_{B \cap I_1} \exp(-\Phi_N(x)) dx = \int_{B \cap I_1} \exp\left(-\frac{1}{2} \sum_{i=1}^{k-1} \left(\lambda_i - \frac{\lambda_i^2}{k}\right) x_i^2 - \frac{k^3}{12} x_k^4 \sum_{i=1}^k V_{ki}^4\right) dx.$$

Secondly, we show that the outer region does not contribute to the limit $N \rightarrow \infty$. It can be seen by elementary tools that $\tilde{\Phi}_N$ has a unique minimum 0 in 0, and so for any $r > 0$ we have $\inf_{x \in I_3} \tilde{\Phi}(x) > 0$. Using the monotone convergence theorem, we obtain

$$\lim_{N \rightarrow \infty} \int_{B \cap I_3} \exp(-N \tilde{\Phi}(x)) dx = 0.$$

Lastly, we will estimate the contribution of the intermediate region from above by a quantity which vanishes as $R \rightarrow \infty$. To this end, we will bound the function Φ_N from below. Recall that

$$\begin{aligned} \Phi_N(x) &= \frac{1}{2} \langle x, \hat{C}_N x \rangle - \frac{N}{2k} \sum_{i=1}^k (V^T \tilde{x})_i^2 + \frac{N}{12k} \sum_{i=1}^k (V^T \tilde{x}_i)^4 + \frac{N}{k} O(\|V^T \tilde{x}\|_6^6) \\ &= \frac{1}{2} \langle x, \hat{C}_N x \rangle - \frac{N}{2k} \langle \tilde{x}, \tilde{x} \rangle + \frac{N}{12k} \|V^T \tilde{x}\|_4^4 + \frac{N}{k} O(\|V^T \tilde{x}\|_6^6) \end{aligned}$$

and since $\|V^T \tilde{x}_i\|_4^4 \geq C \|\tilde{x}_i\|_4^4$ for $C = \|V\|_{4 \rightarrow 4}^{-4}$ this yields

$$\begin{aligned} \Phi_N(x) &\geq \frac{1}{2} \langle x, \hat{C}_N x \rangle - \frac{N}{2k} \langle \tilde{x}, \tilde{x} \rangle + \frac{N}{12k} C \|\tilde{x}\|_4^4 + \frac{N}{k} O(\|V^T \tilde{x}\|_6^6) \\ &= \frac{1}{2} \langle (\Lambda - k^{-1} \Lambda) x, x \rangle + \frac{k^4}{12} C x_k^4 + O(\|V^T \tilde{x}\|_6^6). \end{aligned}$$

Now, as in the case of the central limit theorem, we can estimate from below the error term in such a way that there is a positive constant c and a positive definite matrix C such that

$$\Phi_N(x) \geq \frac{1}{2} \langle C(x_1, \dots, x_{k-1}, 0), (x_1, \dots, x_{k-1}, 0) \rangle + c x_k^4,$$

from which we obtain an upper bound, i.e.

$$\int_{B \cap I_3} \exp(-\Phi_N(x)) dx \leq \int_{B \cap I_3} \exp\left(-\frac{1}{2} \sum_{i,j=1}^{k-1} C_{ij} x_i x_j - c x_k^4\right) dx,$$

and the right hand side vanishes as $R \rightarrow \infty$ by dominated convergence. As a result, the limit $n \rightarrow \infty$ exists and is equal to

$$\lim_{n \rightarrow \infty} \mathbb{P}(w'_n(X_n) + Y_n \in B) = Z^{-1} \int_B \exp\left(-\frac{1}{2} \sum_{i=1}^{k-1} \left(\lambda_i - \frac{\lambda_i^2}{k}\right) x_i^2 - \frac{k^3}{12} x_k^4 \sum_{i=1}^k V_{ki}^4\right) dx.$$

The convergence results for the non-convoluted vector follow easily by considering the characteristic functions. We have for any $t \in \mathbb{R}^k$

$$\mathbb{E} \exp(i \langle t, w'_n(X_n) + Y_n \rangle) \rightarrow \exp\left(-\frac{1}{2} \langle (t_1, \dots, t_{k-1}, \tilde{\Sigma}(t_1, \dots, t_{k-1})) \rangle\right) \phi(t_k),$$

where $\tilde{\Sigma} = \text{diag}(\lambda_i^{-1} + (k - \lambda_i)^{-1})$ and ϕ is the characteristic function of a random variable with distribution $\exp(-x_k^4 k^3 / 12 \sum_{i=1}^k V_{ki}^4)$. Using the independence of X_n and Y_n , the results follow by simple calculations. \square

3.3 Central Limit Theorem: Stein's Method

Lastly, we will prove Theorem 4 using Stein's method of exchangeable pairs. For brevity's sake, for the rest of this section we fix $n \in \mathbb{N}$ and we will drop all sub- and superscripts (e.g. we write B_i instead of $B_i^{(n)}$, \hat{m} instead of $\hat{m}^{(n)}$, J instead of J_n et cetera). It is more convenient to formulate this approach in terms of random variables. Let X be a random vector with distribution μ_J and I be an independent random variable uniformly distributed on $\{1, \dots, N\}$. First, denote by (X, \tilde{X}) the exchangeable pair which is given by taking a step in the Glauber chain for μ_J , i.e. \tilde{X} is the vector after replacing X_I by an independent \tilde{X}_I with distribution $\tilde{X}_I \sim \mu_J(\cdot \mid \bar{X}_I)$ (the exchangeability follows from the reversibility of the Glauber dynamics). Consequently, $(\hat{m}, \hat{m}') = (\hat{m}(X), \hat{m}(\tilde{X}))$ is also exchangeable. More precisely, with the standard basis vectors $(e_i)_{i=1, \dots, k}$ of \mathbb{R}^k we have

$$\hat{m}' := \hat{m} - \frac{X_I - \tilde{X}_I}{\sqrt{|B_I|}} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \Rightarrow \hat{m} - \hat{m}' = \frac{X_I - \tilde{X}_I}{\sqrt{M}} e_{h(I)}. \quad (9)$$

We need the following lemma to identify the conditional expectation of \tilde{X}_i . Here, we write $h : \{1, \dots, N\} \rightarrow \{1, \dots, k\}$ for the function that assigns to each position its block, i.e. $h(j) = k \iff j \in B_k$.

Lemma 4 *Let $\mathcal{F} = \sigma(X)$ and (X, \tilde{X}) be defined as above. Then for each fixed $i \in \{1, \dots, N\}$*

$$\mathbb{E}(\tilde{X}_i \mid \mathcal{F}) = \tanh \left(\frac{1}{\sqrt{N}} (A\Gamma\hat{m})_i - \frac{1}{N} A_{h(i)h(i)} X_i \right).$$

Proof For any Ising model $\mu = \mu_J$ the conditional distribution of \tilde{X}_i is given by $\mu(\cdot \mid \bar{X}_i)$ and so

$$\mathbb{E}(\tilde{X}_i \mid \mathcal{F}) = 2\mu(1 \mid \bar{X}_i) - 1 = \tanh \left((J^{(d)} X)_i \right),$$

where we recall the notation $J^{(d)}$ for the matrix without its diagonal, i.e. $J^{(d)} = J - \text{diag}(J_{ii})$. In the case that $J = J_n$ is the block model matrix, this yields

$$\begin{aligned} \mathbb{E}(\tilde{X}_i \mid \mathcal{F}) &= \tanh \left(N^{-1} \sum_{j=1}^k A_{h(i)j} \sum_{l \in B_j} X_l - N^{-1} A_{h(i)h(i)} X_i \right) \\ &= \tanh \left(N^{-1} (Am)_{h(i)} - N^{-1} A_{h(i)h(i)} X_i \right) \\ &= \tanh \left(N^{-1/2} (A\Gamma\hat{m})_i - N^{-1} A_{h(i)h(i)} X_i \right). \end{aligned}$$

□

Since the conditional expectation will be of importance, we define

$$g_i(X) := N^{-1} (Am)_{h(i)} - N^{-1} A_{h(i)h(i)} X_i = N^{-1/2} (A\Gamma\hat{m})_i - N^{-1} A_{h(i)h(i)} X_i,$$

so that $\mathbb{E}(\tilde{X}_i \mid \mathcal{F}) = \tanh(g_i(X))$. Note that g_i actually does not depend on X_i , the latter term is added for convenience to rewrite the first term. Thus we have $g_i(X) = \mathbb{E}(\tilde{X}_i \mid \bar{X}_i)$.

Lemma 5 *We have*

$$\mathbb{E}(\hat{m} - \hat{m}' \mid \mathcal{F}) = N^{-1} (\text{Id} - \Gamma A \Gamma) \hat{m} + R(X),$$

with

$$R(X) := N^{-1} \sum_{i=1}^k e_i \left((\Gamma A \Gamma \hat{m})_i - |B_i|^{-1/2} \sum_{j \in B_i} \tanh(g_j(X)) \right).$$

Proof From Eq. (9) and Lemma 4 we obtain

$$\begin{aligned} \mathbb{E}(\hat{m} - \hat{m}' \mid \mathcal{F}) &= N^{-1} \sum_{i=1}^k e_i |B_i|^{-1/2} \sum_{j \in B_i} \mathbb{E}(X_j - \widetilde{X}_j \mid \mathcal{F}) \\ &= N^{-1} \sum_{i=1}^k e_i \hat{m}_i - N^{-1} \sum_{i=1}^k e_i |B_i|^{-1/2} \sum_{j \in B_i} \tanh(g_j(X)) \\ &= N^{-1} \hat{m} - N^{-1} \sum_{i=1}^k e_i |B_i|^{-1/2} \left(\sum_{j \in B_i} N^{-1/2} (A \Gamma \hat{m})_i \right) + R(X) \\ &= N^{-1} (\text{Id} - \Gamma A \Gamma) \hat{m} + R(X). \end{aligned}$$

□

For n large enough, the matrix $\Lambda := N^{-1}(\text{Id} - \Gamma A \Gamma)$ satisfies $\|\Lambda\|_{2 \rightarrow 2} < \frac{1}{N}$ and is thus invertible, with inverse $\Lambda^{-1} = N \sum_{l=0}^{\infty} (\Gamma A \Gamma)^l$. Moreover, we also have $\|\Lambda^{-1}\|_{2 \rightarrow 2} \leq N(1 - \|\Gamma A \Gamma\|_{2 \rightarrow 2})^{-1}$.

We will need the following approximation theorem for random vectors.

Theorem 5 ([30, Theorem 2.1]) *Assume that (W, W') is an exchangeable pair of \mathbb{R}^d -valued random vectors such that*

$$\mathbb{E} W = 0, \quad \mathbb{E} W W^t = \Sigma,$$

with $\Sigma \in \mathbb{R}^{d \times d}$ symmetric and positive definite. Suppose further that

$$\mathbb{E}[W' - W \mid W] = -\Lambda W + R$$

is satisfied for an invertible matrix Λ and a $\sigma(W)$ -measurable random vector R . Then, if Z has d -dimensional standard normal distribution, we have for every three times differentiable function

$$|\mathbb{E} h(W) - \mathbb{E} h(\Sigma^{1/2} Z)| \leq \frac{|h|_2}{4} E_1 + \frac{|h|_3}{12} E_2 + \left(|h|_1 + \frac{1}{2} d \|\Sigma\|^{1/2} |h|_2 \right) E_3,$$

where, with $\lambda(i) := \sum_{m=1}^d |(\Lambda^{-1})_{m,i}|$, we define the three error terms

$$\begin{aligned} E_1 &= \sum_{i,j=1}^d \lambda(i) \sqrt{\text{Var} \mathbb{E}[(W'_i - W_i)(W'_j - W_j) \mid W]}, \\ E_2 &= \sum_{i,j,k=1}^d \lambda(i) \mathbb{E}|(W'_i - W_i)(W'_j - W_j)(W'_k - W_k)|, \\ E_3 &= \sum_{i=1}^d \lambda(i) \sqrt{\text{Var} R_i}. \end{aligned}$$

Here, $|h|_j$ denotes the supremum of the partial derivatives of up to order j .

Note that in the proof the choice of $\sigma(W)$ for the conditional expectation is arbitrary; it suffices to take any σ -algebra \mathcal{F} with respect to which W is measurable. Clearly, the value E_1 has to be adjusted accordingly.

Corollary 2 *Let \hat{m} be the block magnetization vector and \hat{m}' as above, define $\Sigma := \mathbb{E} \hat{m} \hat{m}^T$ and let $Z \sim \mathcal{N}(0, \Sigma)$. For any function $h \in \mathcal{F}_3$*

$$|\mathbb{E} h(\hat{m}(X)) - \mathbb{E} h(Z)| \leq CN \left(\frac{|h|_2}{4} E_1 + \frac{|h|_3}{12} E_2 + \left(|h|_1 + \frac{1}{2} k \|\Sigma\|^{1/2} |h|_2 \right) E_3 \right)$$

with the three error terms

$$\begin{aligned} E_1 &= \sum_{i=1}^k \sqrt{\text{Var}(\mathbb{E}((\hat{m}_i(X) - \hat{m}_i(\tilde{X}))^2 | \mathcal{F}))} \\ E_2 &= \sum_{i=1}^k \mathbb{E} |\hat{m}_i(X) - \hat{m}_i(\tilde{X})|^3 \\ E_3 &= \sum_{i=1}^k \sqrt{\text{Var}(R_i)}. \end{aligned}$$

Finally, the following lemma shows that all error terms E_i can be bounded by a term of order $N^{-3/2}$.

Lemma 6 *In the situation of Corollary 2 we have*

$$\max(E_1, E_2, E_3) = O(N^{-3/2}).$$

Before we prove this lemma (and consequently Theorem 4), we will state concentration of measure results in the block spin Ising models. These will be necessary to bound E_1, E_2, E_3 . The first step is the existence of a logarithmic Sobolev inequality for the Ising model μ_{J_n} with a constant that is uniform in n .

Proposition 3 *Under the general assumptions, if $\|\Gamma_\infty A \Gamma_\infty\|_{2 \rightarrow 2} < 1$, then for n large enough the Ising model μ_{J_n} satisfies a logarithmic Sobolev inequality with a constant $\sigma^2 = \sigma^2(\|\Gamma_\infty A \Gamma_\infty\|_{2 \rightarrow 2})$, i.e. for any function $f : \{-1, +1\}^N \rightarrow \mathbb{R}$ we have*

$$\text{Ent}_{\mu_{J_n}}(f^2) \leq 2\sigma^2 \sum_{i=1}^N \mathbb{E}_{\mu_{J_n}}(f - f \circ T_i)^2, \quad (10)$$

where Ent is the entropy functional and $T_i(\sigma) = (\sigma_1, \dots, \sigma_{i-1}, -\sigma_i, \sigma_{i+1}, \dots, \sigma_N)$ the sign flip operator.

This follows immediately from [23, Proposition 1.1], since $\Gamma_n A \Gamma_n \rightarrow \Gamma_\infty A \Gamma_\infty$, which implies the convergence of the norms, i.e. for n large enough we have $\|\Gamma_n A \Gamma_n\|_{2 \rightarrow 2} < 1$. Although the condition in [23] is $\|J\|_{1 \rightarrow 1} < 1$, this was merely for applications' sake and $\|J\|_{2 \rightarrow 2} < 1$ is sufficient to establish the logarithmic Sobolev inequality.

For any function $f : \{-1, +1\}^N \rightarrow \mathbb{R}$ and any $r \in \{1, \dots, N\}$ we write

$$\mathfrak{h}_r f(x) = |f(x) - f(T_r x)|,$$

so that (10) becomes

$$\text{Ent}_{\mu_{J_n}}(f^2) \leq 2\sigma^2 \sum_{r=1}^N \int (\mathfrak{h}_r f(x))^2 d\mu_{J_n}(x).$$

Moreover, it is known that (10) implies a Poincaré inequality

$$\text{Var}(f) \leq \sigma^2 \sum_{r=1}^N \mathbb{E} \mathfrak{h}_r f(X)^2. \quad (11)$$

Proof of Lemma 6 Error term E_1 : To treat the term E_1 , fix $i \in \{1, \dots, k\}$ and observe that

$$\begin{aligned} \mathbb{E}((\hat{m}_i(X) - \hat{m}_i(\tilde{X}))^2 \mid \mathcal{F}) &= N^{-1} \sum_{j=1}^N \mathbb{E}((\hat{m}_i(X) - \hat{m}_i(\bar{X}_j, \tilde{X}_j))^2 \mid \mathcal{F}) \\ &= (N|B_i|)^{-1} \sum_{j \in B_i} \mathbb{E}((X_j - \tilde{X}_j)^2 \mid \mathcal{F}) \\ &= -2(N|B_i|)^{-1} \sum_{j \in B_i} X_j \tanh(g_j(X)) + 2N^{-1}. \end{aligned}$$

Thus, if we define

$$f_i(X) := |B_i^{(n)}|^{-1/2} \sum_{j \in B_i} X_j \tanh\left(N^{-1} \sum_{l=1}^k A_{il} m_l(X) - N^{-1} A_{ii} X_i\right),$$

we see that

$$\text{Var}^{1/2}\left(\mathbb{E}\left((\hat{m}_i^{(n)} - \hat{m}_i^{(n)})^2 \mid \mathcal{F}\right)\right) = 2N^{-1}|B_i^{(n)}|^{-1/2} \text{Var}^{1/2}(f_i(X)),$$

and we need to show that $\text{Var}(f_i(X)) = O(1)$. Using the Poincaré inequality (11) it suffices to prove that $\mathfrak{h}_r f_i(X)^2 \leq C|B_i^{(n)}|^{-1}$.

Let $r \in \{1, \dots, N\}$ be arbitrary and define $h_i(X) := N^{-1} \sum_{l=1}^k A_{il} m_l(X) - N^{-1} A_{ii} X_i$. The first case is that $r \in B_i^{(n)}$, for which

$$\begin{aligned} \mathfrak{h}_r f_i(X) &\leq |B_i|^{-1/2} |2X_r \tanh(h_i(X))| + |B_i|^{-1/2} \sum_{\substack{j \in B_i \\ j \neq r}} |\tanh(h_i(X)) - \tanh(h_i(T_r X))| \\ &\leq 4|B_i|^{-1/2} + |B_i|^{-1/2} N^{-1} \sum_{\substack{j \in B_i \\ j \neq r}} |*| \sum_{l=1}^k A_{il} (m_l(X) - m_l(T_r(X))) \\ &\leq |B_i|^{-1/2} (4 + 2\|A\|_\infty). \end{aligned}$$

The second case $r \notin B_i^{(n)}$ follows by similar reasoning.

Error term E_2 : The second term E_2 is much easier to estimate, as

$$\mathbb{E}|\hat{m}_i - \hat{m}_i'|^3 = N^{-1}|B_i|^{-3/2} \sum_{j \in B_i} \mathbb{E}|X_j - \tilde{X}_j|^3 \leq 8N^{-1}|B_i|^{-1/2} = O(N^{-3/2}).$$

Error term E_3 : To estimate the variance of the remainder term R we first split it into two sums. For any $i = 1, \dots, k$ write

$$\begin{aligned} R_i(X) &= N^{-1} \left(|B_i|^{-1/2} \sum_{j \in B_i} g_j(X) - \tanh(g_j(X)) + N^{-1} A_{ii} X_j \right) \\ &= N^{-1} |B_i|^{-1/2} \sum_{j \in B_i} g_j(X) - \tanh(g_j(X)) + N^{-2} A_{ii} m_i(X) \\ &=: R_j^{(1)}(X) + R_j^{(2)}(X). \end{aligned}$$

Clearly $\|R_i - \mathbb{E} R_i\|_2 \leq \|R_i^{(1)} - \mathbb{E} R_i^{(1)}\|_2 + \|R_i^{(2)} - \mathbb{E} R_i^{(2)}\|_2$ and we estimate these terms separately. It is obvious that the L^2 norm of the second term is of order $O(N^{-2})$. To estimate $R_i^{(1)}$, we use $\tanh(x) - x = O(x^3)$ to obtain

$$\begin{aligned} \|R_i^{(1)} - \mathbb{E} R_i^{(1)}\|_2 &\leq C N^{-1} |B_i|^{-1/2} \sum_{j \in B_i} \| |g_j(X)|^3 \|_2 \\ &\leq C N^{-1} |B_i|^{-1/2} \sum_{j \in B_i} \| |N^{-1/2} (A\Gamma \hat{m})_j|^3 \|_2 + \| N^{-3} |A_{ii}|^3 \|_2 \\ &= O(N^{-2}) + O(N^{-5/2}). \end{aligned}$$

In the last line we have used the fact that $\|(A\Gamma \hat{m})_i^3\|_2 = \|A\Gamma \hat{m}_i\|_6^3$ and for all $p \geq 2$

$$\|(A\Gamma \hat{m})_i\|_p \leq C \sum_{l=1}^k \|\hat{m}_l\|_p \leq C \sum_{l=1}^k (\sigma^2 p)^{1/2}$$

which evaluated at $p = 6$ gives $\|(A\Gamma \hat{m})_i\|_6^3 = O(1)$. For the details see [23]. The constant depends on a norm of $A\Gamma$, which by convergence to $A\Gamma_\infty$ can again be chosen independently of n . \square

Proof of Theorem 4 The theorem follows immediately from Corollary 2 and Lemma 6. \square

4 Discussion and Open Questions

Although the questions raised in the introduction have been answered to a certain degree, there are still open questions that we were not yet able to answer.

The first question concerns the maxima of the rate function I . Firstly, note that by [11, Theorem A.1] the global maxima of I are related to the global minima of the so-called *pressure functional*, which can for example be found in [17, equation (14)]. Using the compactness of $[-1, 1]^k$ and the continuity of I , the existence of a maximiser easily follows, but the number of maximisers is still obscure. From real-analyticity of I , we can infer that the set of maximisers is a λ^k null set, but it could in principle contain infinitely many points. However, Lemmas 2 and 3 as well as numerics suggest that for positive interactions and $k \geq 2$, the number of local minima is twice the number of independent systems - see Figures 2 for the $k = 3$ and 3 for the $k = 2$ case below.

However, we believe that the case of negative interactions between groups might drastically change the picture. Indeed, consider a model with three blocks and positive interaction β within the blocks and negative interaction α between the blocks. Then, if β is large enough, the points within the blocks will tend to be aligned. However, as α is negative, the magnetizations

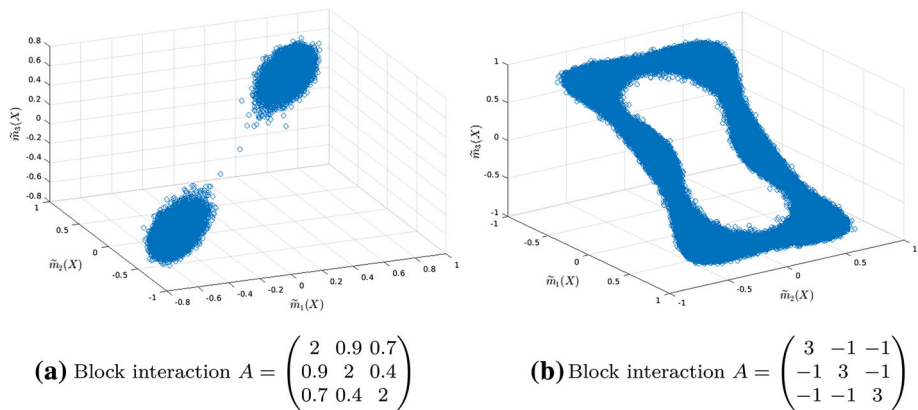


Fig. 2 A scatterplot of the normalized block magnetization \tilde{m} in the uniform case with $k = 3$ blocks and $n = 500$, sampled using the Glauber dynamics—note that it is not rapidly mixing!

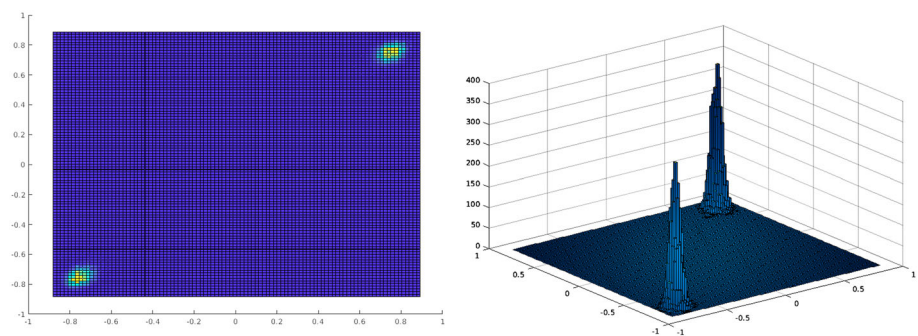


Fig. 3 A heatmap (left) and a histogram (right) of the block magnetization vector $m = (m^1, m^2)$ in the uniform, low temperature case. The block interaction matrix is given by $A = \begin{pmatrix} 1.8 & 0.8 \\ 0.8 & 1.8 \end{pmatrix}$

of block one and two will try to have different signs, but so do the magnetizations of blocks two and three, and three and one. Hence, frustration occurs. In this respect, a model with positive and negative interactions carries features of a spin glass.

Another question is the relationship of Theorems 2 and 4. In Theorem 4 we consider the distance to a normal distribution with covariance matrix $\Sigma_n := \mathbb{E} \hat{m}^{(n)} (\hat{m}^{(n)})^T$ and not to $\Sigma_\infty := (\text{Id} - \Gamma_\infty A \Gamma_\infty)^{-1}$, which is the covariance matrix of the limiting distribution. Testing against functions $h \in C_c^\infty(\mathbb{R}^k)$, we see that Σ_∞ is the limit of the matrices Σ_n . It is an interesting task to provide suitable bounds of $\|\Sigma_n - \Sigma_\infty\|$ in any matrix norm, since [30, Proposition 2.8] provides bounds of $|\mathbb{E} h(X) - \mathbb{E} h(Y)|$ for two random vectors with $X \sim \mathcal{N}(0, \Sigma_0)$ and $Y \sim \mathcal{N}(0, \Sigma_1)$ in terms of the 1-distance of Σ_0 and Σ_1 .

Thirdly, it remains an open problem to quantify the distance to a normal distribution with the “limiting” covariance matrix Σ_∞ . The central limit theorem in the one-dimensional Curie–Weiss model has been solved for example in [14, Corollary 2.9]. Therein one can see that the limiting covariance is $(1 - \beta)^{-1}$ by considering the approximate linear regression condition. A similar condition is true in the multidimensional case. For example, in Lemma 5 we have proven

$$\mathbb{E} \left(\hat{m}^{(n)} - \hat{m}^{(n)'} \mid \mathcal{F} \right) = \lambda \Lambda^{-1} \hat{m}^{(n)} + R(X), \quad (12)$$

where $\lambda = N^{-1}$ and $\Lambda = (\text{Id} - \Gamma_n A \Gamma_n)^{-1}$. Thus, in the case $\Gamma_n \equiv \Gamma_\infty$ (e.g. consider a subsequence along which this holds) Λ is the covariance matrix of the limit distribution. However, we have been unable to find a suitable modification of [30, Theorem 2.1] that enables one to compare the distribution of the random vector $\hat{m}^{(n)}$ with $\mathcal{N}(0, \Lambda)$.

Acknowledgements Open access funding provided by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Agliari, E., Burioni, R., Contucci, P.: A diffusive strategic dynamics for social systems. *J. Stat. Phys.* **139**(3), 478–491 (2010). <https://doi.org/10.1007/s10955-010-9948-1>
2. Amini, A.A., Levina, E.: On semidefinite relaxations for the block model. *Ann. Stat.* **46**(1), 149–179 (2018). <https://doi.org/10.1214/17-AOS1545>
3. Berthet, Q., Rigollet, P., Srivastava, P.: Exact recovery in the Ising blockmodel. *Ann. Stat.* **47**(4), 1805–1834 (2019). <https://doi.org/10.1214/17-AOS1620>
4. Bresler, G.: Efficiently learning Ising models on arbitrary graphs [extended abstract]. In: STOC'15—Proceedings of the 2015 ACM Symposium on Theory of Computing, pp. 771–782. ACM, New York (2015)
5. Bresler, G., Mossel, E., Sly, A.: Reconstruction of Markov random fields from samples: some observations and algorithms. *SIAM J. Comput.* **42**(2), 563–578 (2013). <https://doi.org/10.1137/100796029>
6. Brock, W.A., Durlauf, S.N.: Discrete choice with social interactions. *Rev. Econom. Stud.* **68**(2), 235–260 (2001). <https://doi.org/10.1111/1467-937X.00168>
7. Chatterjee, S., Shao, Q.M.: Nonnormal approximation by Stein's method of exchangeable pairs with application to the Curie–Weiss model. *Ann. Appl. Probab.* **21**(2), 464–483 (2011). <https://doi.org/10.1214/10-AAP712>
8. Collet, F.: Macroscopic limit of a bipartite Curie–Weiss model: a dynamical approach. *J. Stat. Phys.* **157**(6), 1301–1319 (2014). <https://doi.org/10.1007/s10955-014-1105-9>
9. Comets, F.: Large deviation estimates for a conditional probability distribution. Applications to random interaction Gibbs measures. *Probab. Theory Relat. Fields* **80**(3), 407–432 (1989). <https://doi.org/10.1007/BF01794432>
10. Cont, R., Löwe, M.: Social distance, heterogeneity and social interactions. *J. Math. Econom.* **46**(4), 572–590 (2010). <https://doi.org/10.1016/j.jmateco.2010.03.009>
11. Costeniu, M., Ellis, R.S., Touchette, H.: Complete analysis of phase transitions and ensemble equivalence for the Curie–Weiss–Potts model. *J. Math. Phys.* **46**(6), 063301 (2005). <https://doi.org/10.1063/1.1904507>
12. Dembo, A., Zeitouni, O.: Large Deviations Techniques and Applications. Stochastic Modelling and Applied Probability, vol. 38. Springer, Berlin (2010)
13. den Hollander, F.: Large Deviations, Fields Institute Monographs, vol. 14. American Mathematical Society, Providence (2000)
14. Eichelsbacher, P., Löwe, M.: Stein's method for dependent random variables occurring in statistical mechanics. *Electron. J. Probab.* **15**(30), 962–988 (2010). <https://doi.org/10.1214/EJP.v15-777>
15. Ellis, R.S.: Entropy, Large Deviations, and Statistical Mechanics. Classics in Mathematics. Springer, Berlin (2006)
16. Ellis, R.S., Newman, C.M.: Limit theorems for sums of dependent random variables occurring in statistical mechanics. *Z. Wahrsch. Verw. Gebiete* **44**(2), 117–139 (1978). <https://doi.org/10.1007/BF00533049>
17. Fedele, M., Contucci, P.: Scaling limits for multi-species statistical mechanics mean-field models. *J. Stat. Phys.* **144**(6), 1186–1205 (2011). <https://doi.org/10.1007/s10955-011-0334-4>

18. Fedele, M., Unguendoli, F.: Rigorous results on the bipartite mean-field model. *J. Phys. A* **45**(38), 385001, 18 (2012). <https://doi.org/10.1088/1751-8113/45/38/385001>
19. Gallo, I., Contucci, P.: Bipartite mean field spin systems. Existence and solution. *Math. Phys. Electron. J.* **14**, 1, 21 (2008)
20. Gallo, I., Barra, A., Contucci, P.: Parameter evaluation of a simple mean-field model of social interaction. *Math. Models Methods Appl. Sci.* **19**(suppl.), 1427–1439 (2009). <https://doi.org/10.1142/S0218202509003863>
21. Gao, C., Ma, Z., Zhang, A.Y., Zhou, H.H.: Achieving optimal misclassification proportion in stochastic block models. *J. Mach. Learn. Res.* **18**, 60, 45 (2017)
22. Gentz, B., Löwe, M.: The fluctuations of the overlap in the Hopfield model with finitely many patterns at the critical temperature. *Probab. Theory Relat. Fields* **115**(3), 357–381 (1999). <https://doi.org/10.1007/s004400050241>
23. Götze, F., Sambale, H., Sinulis, A.: Higher order concentration for functions of weakly dependent random variables. *Electron. J. Probab.* **24**(85), 19 (2019). <https://doi.org/10.1214/19-EJP338>
24. Kincaid, J.M., Cohen, E.G.D.: Phase diagrams of liquid helium mixtures and metamagnets: experiment and mean field theory. *Phys. Rep.* **22**(2), 57–143 (1975). [https://doi.org/10.1016/0370-1573\(75\)90005-8](https://doi.org/10.1016/0370-1573(75)90005-8)
25. Kirsch, W., Toth, G.: Two groups in a Curie-Weiss model with heterogeneous coupling. *J. Theor. Probab.* (2019). <https://doi.org/10.1007/s10959-019-00933-w>
26. Knöpfel, H., Löwe, M.: Zur Meinungsbildung in einer heterogenen Bevölkerung—ein neuer Zugang zum Hopfield Modell. *Math. Semesterber.* **56**(1), 15–38 (2009). <https://doi.org/10.1007/s00591-008-0049-z>
27. Löwe, M., Schubert, K.: Fluctuations for block spin Ising models. *Electron. Commun. Probab.* **23**, 53, 12 (2018). <https://doi.org/10.1214/18-ECP161>
28. Mossel, E., Neeman, J., Sly, A.: Belief propagation, robust reconstruction and optimal recovery of block models. *Ann. Appl. Probab.* **26**(4), 2211–2256 (2016). <https://doi.org/10.1214/15-AAP1145>
29. Opoku, A.A., Owusu Edusei, K., Ansah, R.K.: A conditional Curie–Weiss model for stylized multi-group binary choice with social interaction. *J. Stat. Phys.* **171**(1), 106–126 (2018). <https://doi.org/10.1007/s10955-018-1988-y>
30. Reinert, G., Röllin, A.: Multivariate normal approximation with stein’s method of exchangeable pairs under a general linearity condition. *Ann. Probab.* **37**(6), 2150–2173 (2009). <https://doi.org/10.1214/09-AOP467>
31. van Hemmen, J.L., van Enter, A.C.D., Canisius, J.: On a classical spin glass model. *Z. Phys. B* **50**(4), 311–336 (1983). <https://doi.org/10.1007/BF01470043>
32. van Hemmen, J.L., Gensing, D., Huber, A., Kühn, R.: Elementary solution of classical spin-glass models. *Z. Phys. B* **65**(1), 53–63 (1986). <https://doi.org/10.1007/BF01308399>

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.